Login

SAP Predictive Analytics Tutorial: Music Recommendation

August 10, 2016 | 2,552 Views |

Andreas Forster
more by this author

SAP Predictive Analytics

Retail | analytics | former jive document | predictive analytics | retail | tutorial-

G+ share   f share   🐦 tweet   in share   like   Follow

This tutorial shows how to make a shopping basket analysis to create product recommendations. The "Social Network Analysis" functionality from the automated mode of SAP Predictive Analytics can easily find such product suggestions.

You see such recommendations for instance in online shops, which point out complementary products. Typically, there are way too many options for someone to create such rules manually. Instead such recommendations are derived from behaviour seen by previous customers. With this tutorial you can create such suggestions yourself, even if you are not a Data Scientist.

To make it more fun I am using a dataset with people's listening history to find music recommendations. Who knows, you might find some new music you enjoy along the way....

The musical history I am using is coming from a very cool project called #nowplaying, which is monitoring social media for posts about music people are currently listening. That information is extracted, cleaned, collected and shared on their website. As of August 2016 they have about 50 million tweets! I would like to thank Eva Zangerle and Martin Pichl who are behind that project, for their effort and for making their data available!

Dataset

For this tutorial I have taken the data file from #nowplaing and filtered on the last 2 years (from July 2014 to June 2016). Aggregating the data by user and artist tells us how often each user has listened to any band in that period. That is the history we will derive the recommendations from. To make sure people really enjoyed the music, another filter was applied to remove any combinations when a person listened less than 5 times to a band in two years. In that case they don't seem to be very keen on that music. On my computer some band names could not be displayed correctly. simply deleted those. (see the following chapter for the download link of the file used in this tutorial)

This has left about 700.000 combinations of 68.000 users and 76.000 bands. Here is an extract

USER	ARTIST	LISTENCOUNT
39c364a8d7822a0a605d7e9dc79d5d864a43034f	Coldplay	10
e0bf7222ed152b5f059f03f44da3bf764118f65b	Avril Lavigne	8
2f1ab8a9ff62614db8a65d1fb687eb30efb80253	Gorillaz	146
39c364a8d7822a0a605d7e9dc79d5d864a43034f	Ellie Goulding	11
15aff30694d1f668baae6a0d6bcd92fa19a0df67	2Pac	25

You get the idea. The user name's have been made anonymous with a unique has value. Nobody's taste in music is being judged….

# Setup

To follow the tutorial you just need two things:

- SAP Predictive Analytics: If you don't own a copy, you can use a 30-day trial license currently available free of charge here on SCN.
- And the data of course. Download NOWPLAYINGLISTENCOUNT5PLUS.txt from GitHub.

Start the Analysis

Start "SAP Predictive Analytics". On the menu on the left go into "Social". In the center click on "Create a Social Network Analysis". The term "Network" might be confusing. This is not really referring to a network like Facebook or an IT network. In our case it refers to a network of users and bands/artists (or customers and products). We have lots of users and lots of bands. A user is connected to a band if he has listened to their music. This results in a huge network, which is the basis for the recommendations.
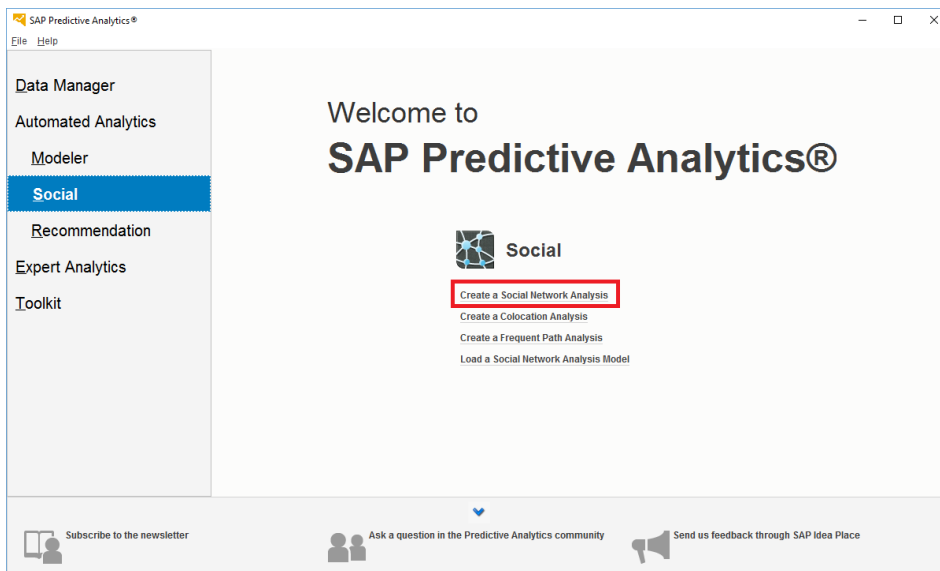
Select "Build a Social Graph From a Data Set" and choose the file NOWPLAYINGLISTENCOUNT5PLUS.txt (the download link is above).

Continue with "Next". Hit the "Analyze"-icon and the tool will pick up the column headers and types. The column "KxIndex" does not exist in the file. It was added by the tool automatically. We will not be using it however.

Just "OK" that screen and you can already define the initial network, which is called here a graph. This graph connects the users and artists, so make the following settings:

Graph Name: NowPlaying

Graph Creation Type: Transactions

Source Node: USER

Target Node: ARTIST

Use a weight column: LISTENCOUNT

Transactions can be thought of here as products purchases or bands listened to. In a supermarket this would be lists of products purchased by

a person. Here it is the artists the person listened to. So the same concept would apply to point of sale data from a retailer. The weight, which gives an indication of the strength of the link, might then be the quantity or the price.

Your screen should look like this:

Do not click anywhere on the bottom right…. Whilst on this screen, click the "Add Graph" icon on the top left. Then select "Derive Graph From a Bipartite Graph". So far in the earlier steps we have created a so called bipartite graph. This means the graph contains 2 parties: users and artists. Out of that graph, we now create a second graph of artists only, based on how they are connected through their listeners. If one user has listened to two bands in the bipartite graph, then these bands are connected in the derived graph.

You just need to change the Entity to "ARTIST".

Now "OK" this and click "Next" until you see a summary before the calculation starts. Don't worry about the settings that could have been changed along the way.
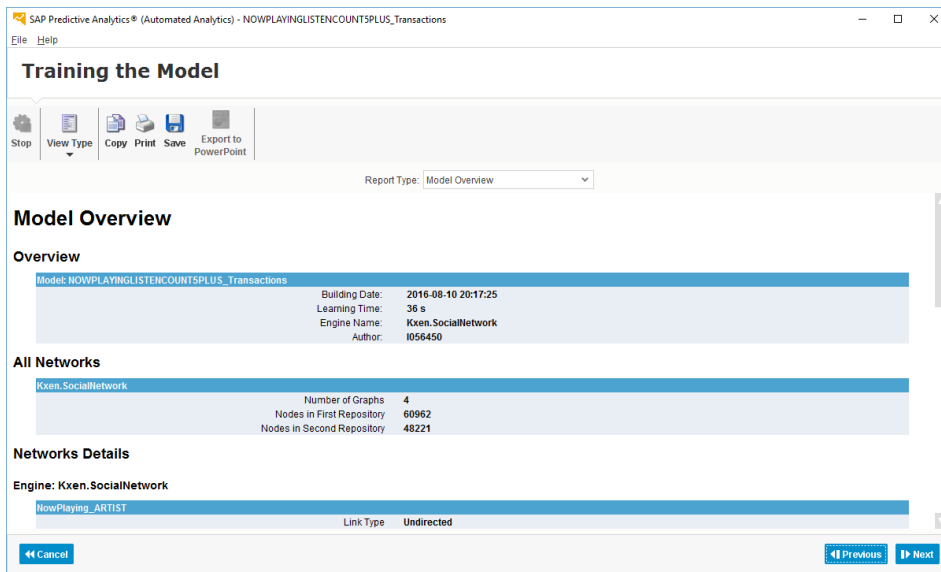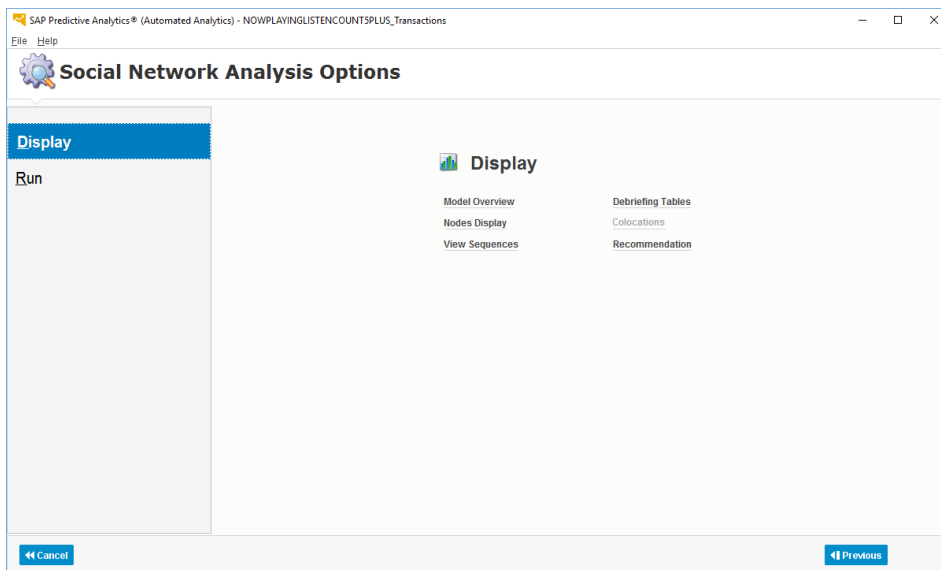

Finally click "Generate". On my laptop the calculation took just over half a minute. Some of the default values restricted the size of the graph. Eventually you can calculate more detailed graphs, which will surely take longer to complete.

Continue to the next screen. Now go into "Nodes Display" to see links between the artists. Each link is a potential recommendation.


On the top change the "Graph" dropdown to "NowPlaying_ARTIST". We are now looking at the 2nd graph that consists only of artists. From the "Node" dropdown select an artist. Don't pick "Alicia Keys" or you drown in recommendations. The weights behind the artist indicate the number of recomendations. You can select "Royal Blood" for instance. You don't have to scroll there, just start typing the name. To see the results, click the "Display Node" icon and select the sub-option with the same name: "Display Node".
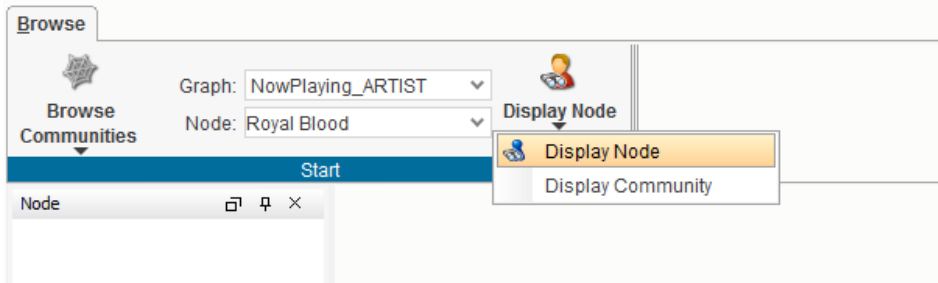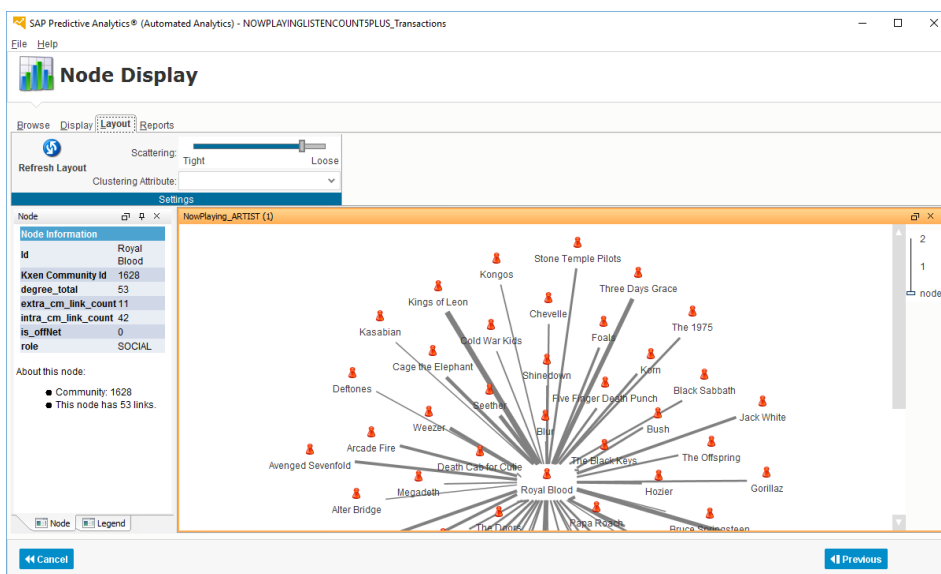
And you see the nodes (other artists) linked to the band you selected!
The bolder the link between two artists, the more often they were
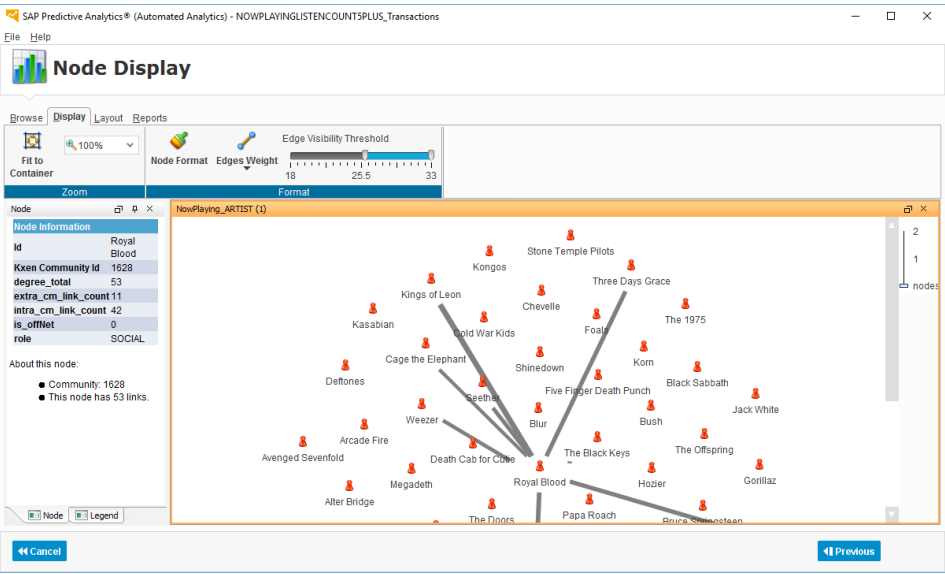listened to by the same users.
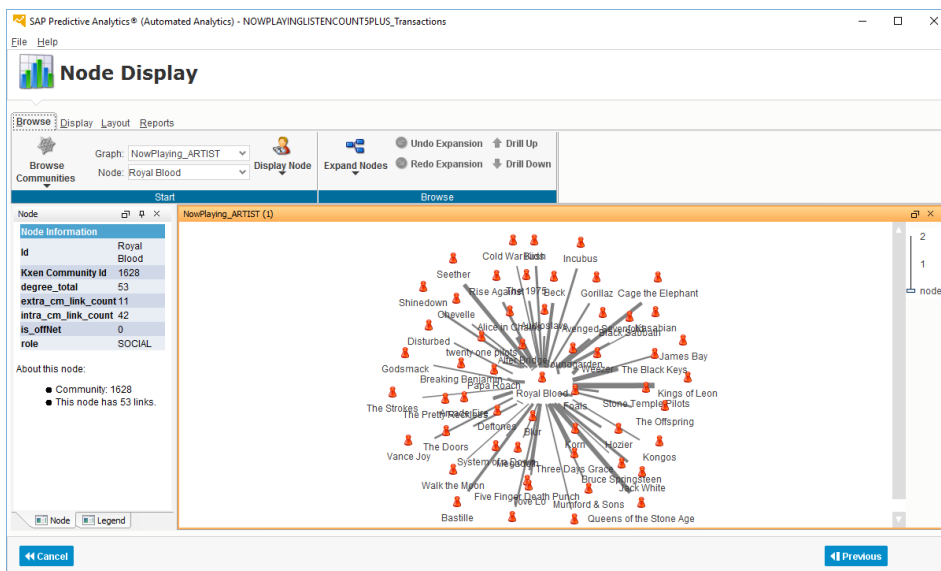
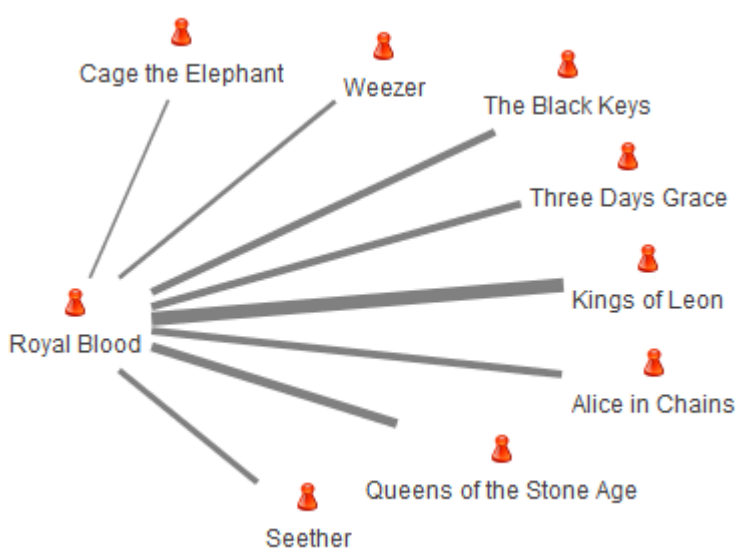On the "Layout" tab you can widen the display.

On the "Display" tab change the slider of the "Edge Visibility Threshold"
to show only the strongest links. We see that if people listen to "Royal

Blood", there is a strong indication they might also enjoy the "Kings of Leon".

Feel free to select orther bands or move the nodes around on the screen.

This has already given good ideas for recommendations. But what about a specific recommendation for a certain user?

Recommendation

Click "Previous" to go to the "Display" screen. Select "Recommendation". Here enter the name/id of the user to create a recommendation for. The user 182cfb7df08c9053b0519122cabbbfc91daf326b for instance did listen to "Royal Blood". Set "Keep Top N" to 20. This will show the 20 most relevant recommendations.

Note that the option "Do not recommend if already owned" is ticked. This means that by default only products will be displayed that have not been bought yet by the user. In our case it means, we will recommend only bands the user has not yet listened to.

Hit "Get Recommendation" and we get our user-specific recommendations based on his and everyone else's histroy.

Here we obtain recommendations for one single user. Obviously the whole process can be automated and such recommendations can be produced en masse or even in real time.

You have created a shopping basket analysis!

However, there is a lot more to get out of this analysis. For instance the tool identified communities (clusters) of similar products. This helps decide which recommendations to suggest. It gives guidance for instance on recommendations, that open up new product areas the person has not purchased before. Using these communities and fine-tuning the current model might be a topic for another article...

2 Comments

You must be Logged on to comment or reply to a post.

Former Member

August 15, 2016 at 6:13 am

Great tutorial, thanks Andreas!

Former Member

February 17, 2017 at 5:30 am

Great

Share & Follow

Privacy

Terms of Use

Legal Disclosure

Copyright

Trademark

Cookie Preferences

Sitemap

Newsletter